On the Stochasticity of Ultimatum Games

Tianxiao Qi, Bin Xu, Jinshan Wu and Nicolaas J. Vriend

ISSN 1473-0278

Working Paper No. 926 April 2021

School of Economics and Finance

Queen Mary
University of London

On the Stochasticity of Ultimatum Games*

Tianxiao Qi[†], Bin Xu[‡], Jinshan Wu[§] and Nicolaas J. Vriend[¶]

April 12, 2021

Abstract

Brenner and Vriend (2006) argued (experimentally and theoretically) that one should not expect proposers in ultimatum games to learn to converge to the subgame perfect Nash equilibrium offer, as finding the optimal offer is a hard learning problem for (boundedly-rational) proposers. In this paper we show that providing the proposers with given (fixed) acceptance probabilities (essentially eliminating the learning task) leads to somewhat lower offers, but still substantially above the monetary payoff-maximizing offer. By using a Risk Attitude test and a Probability Matching test, we show experimentally that the proposers' attitude with respect to risk, as well as their ability to interpret and deal with probabilities may matter when it comes to making UG offers. Thus, we argue that the lack of convergence to the minimum offers in ultimatum games may be related to the inherent stochasticity of typical UG experiments, highlighting a possible cause of such deviations that seems a complementary explanation to existing ones.

JEL Classification: C72, C73

Keywords: Ultimatum game, Stochasticity, Risk Attitude, Probability Matching

*We thank Nobi Hanaki, Friederike Mengel and participants of the ASFEE (Nice), WEHIA (London) and ESA (Vancouver) conferences as well as a GREDEG (Nice) seminar for comments and suggestions. This work was partially supported by the National Natural Science Foundation of China (grant number 71973128), Zhejiang Natural Science Foundation Project (No. LY18G030012) and Funded by the Humanities and Social Sciences Research Program of the Ministry of Education (No. 18YJA790092). All errors are our own.

[†]School of Systems Science, Beijing Normal University, Beijing 100875, P.R. China

[‡]School of Economics, Zhejiang Gongshang University, and Modern business research Center, Zhejiang Gongshang University, Hangzhou, 310018, China

[§]Corresponding author: School of Systems Science, Beijing Normal University, Beijing 100875, P.R. China, <jinshanw@bnu.edu.cn>

[¶]Corresponding author: School of Economics and Finance, Queen Mary University of London, <n.vriend@qmul.ac.uk>

1 Introduction

The Ultimatum Game (UG) can be seen as a stylized representation of many social situations in which take-it-or-leave-it offers are made. Ultimatum games have been extensively studied, and their game-theoretic analysis seems straightforward: The proposer offers the smallest possible amount to the receiver, and this offer is accepted.[1] But intriguingly, abundant empirical evidence seems to be at odds with the game-theoretic analysis. Average offers are typically in the 40-50% range, lower offers are often rejected, and in multi-period setups we typically see convergence of offers to levels just below 50% (see e.g. Camerer, 2003). The main explanations offered in the literature seem to fall into two broad categories. First, many explanations relate to fairness considerations in some broad sense, including the effects of emotions and other-regarding preferences in many shapes and forms. Second, other explanations relate to learning, in particular the relative speed of learning (and the relative incentives to learn) of proposers and receivers, and the fact that this is a coevolutionary process, in which the learning of one side has implications for the learning task of the other side.

In this paper we will focus on the behavior of the proposers in ultimatum games. This is not because we think the behavior of the receivers is unimportant or fully understood, but because it allows us to highlight some insights concerning the behavior of proposers. We believe these insights do not only shine some light on the empirical evidence of ultimatum games, but may also be of some more general relevance.

In following this approach we build on Brenner and Vriend (2006). They used an experimental design in which proposers played against artificial receivers (computers). This allowed them to minimize the possible effect of fairness concerns and other-regarding preferences, as well as to abstract from coevolutionary learning processes. Their main conclusion was that finding the optimal offer is a hard learning problem for proposers in ultimatum games, and that the lack of convergence to the optimal offer is an inherent feature of the learning task faced by (boundedly-rational) proposers in these games.

We extend this line of research by further stripping down the set of possible

[1] More precisely, if the strategy space is continuous the subgame perfect Nash equilibrium is of the form "offer 0 and any offer ≥ 0 will be accepted", while if the strategy space is discrete there may be an additional equilibrium of the form "offer $\epsilon > 0$, and any offer > 0 will be accepted, while 0 would be rejected, where ϵ is the smallest possible offer > 0".

explanations for offers above the subgame perfect Nash equilibrium one. Just like was the case for Brenner and Vriend (2006), the objective of our paper is not to claim that fairness (in some broad sense) or (coevolutionary) learning processes are irrelevant, but rather to indicate other factors causing deviations from the subgame perfect Nash equilibrium in UG experiments, and to provide a quantitative assessment of these factors. Our analysis and main findings can be summarized as follows.

First, with proposers playing again against artificial receivers, we eliminate the learning task of proposers by explicitly giving them the (fixed) acceptance probabilities for all possible offers, and compare this with a treatment in which these probabilities are not given. We observe that even with acceptance probabilities explicitly given, offers remain substantially above the optimal minimum offer, although they are somewhat lower than in the treatment where the acceptance probabilities are not given, and they show a similar slight decline over time. This suggests that there must be additional factors explaining the behavior of the proposers. Second, to test the hypothesis that the players may have some difficulty to interpret and deal with probabilities, all subjects do a Probability Matching test. We find that subjects who fail this test (i.e. those prone to anomalous behavior), offer significantly more in the UG than those who pass the test successfully. Third, to further investigate how the subjects' perception of and attitude with respect to the stochasticity of their environment may matter, all subjects do a standard risk attitude test. We find that seriously risk-averse subjects offer significantly more in the ultimatum game. Fourth, we also find that there are significant differences between male and female subjects. In line with some evidence in the bargaining literature, female subjects offer substantially more in the UG. We find that they are also more likely to fail the Probability Matching test, and that they are more often risk-averse. Controlling for Probability Matching test performance and risk attitude, we find that female subjects tend to make significantly higher UG offers than male subjects.

Our overall conclusion is that a significant part of the observed apparent lack of convergence to the subgame perfect Nash equilibrium in UG experiments may be due to the inherent stochasticity that is typical of such experiments. Besides the difficulty that many players may have to operate in such a stochastic environment, some apparent deviations from the optimal minimum offer may be explained by the players' risk attitude. This stochasticity explanation seems complementary to alternative explanations in the literature. We conjecture that

UG experiments are not the only experiments for which it is the case that this sense of stochasticity plays a major role in determining the behavior of the players.

The remainder of this paper is organized as follows. In Section 2 we discuss some related literature. Section 3 presents details of our experimental design, while Section 4 presents the experimental analysis. Section 5 concludes.

2 Some related literature

2.1 Literature: Ultimatum Game (UG)

Since the seminal paper Güth et al. (1982) introduced the Ultimatum Game (UG) to the economics audience, the impact on the economics (experimental and theoretical) profession has been huge, and is well-illustrated by the survey in Güth and Kocher (2014).

In the early years of the UG literature, two main competing explanations appeared concerning the widespread observation that behavior of both proposers and receivers tended to be at odds with the game-theoretic solution based on monetary payoff maximization. A first explanation was related to (perceived) fairness concerns and the possibility for rejections to be used as punishment (see e.g. Güth et al., 1982; Thaler, 1988; Forsythe et al., 1994; Bolton and Zwick, 1995). Some alternative explanations focused on the relative speed of learning of proposers and receivers and how these may relate to (off-equilibrium) incentives (see e.g. Gale et al., 1995; Roth and Erev, 1995).[2]

In subsequent years, it is probably fair to say that attention as far as explanations were concerned was predominantly focused on fairness in some rather broad sense (see e.g. Güth and Kocher, 2014; Cooper and Kagel, 2016).[3] This broad category of fairness considerations encompasses various kinds of other-regarding preferences as well as emotions, and fairness considerations as such have also been related to socio-demographic and cultural factors (see e.g. Henrich, 2000; Oosterbeek et al., 2004; Carpenter et al., 2005; Chuah et al., 2007).

[2]Note, e.g., that for offers close to the minimum, the incentives for the receivers to accept are relatively small while the proposers would lose a lot from a rejection. As a result, proposers may learn to move away from the minimum before receivers have learned to accept such low offers.

[3]Although with e.g. Cooper et al. (2003) there was still some attention for the relative speed of learning type of explanation as well.

One of these socio-demographic factors is gender. The role of gender in bargaining in general and ultimatum bargaining more specifically has been extensively studied. While various gender stereotypes frequently show up, the overall picture seems somewhat mixed. To some extent this seems due to the following two aspects. First, even with bilateral bargaining, a good number of different gender configurations is possible, and even more so if information thereof is a treatment variable. Second, fine details of the bargaining setup and framing may matter when it comes to gender effects. In ultimatum games, Eckel and Grossman (2001) find that women make higher offers. Solnick (2001) does not find a difference in offers, although she confirms that this is what receivers appear to expect, whereas in a replication study (Li et al., 2018) no significant gender differences are found. Garcia-Gallego et al. (2012) analyze an ultimatum game framed as salary bargaining, where the salary relates to either a fictitious or real task. They find no gender difference for the former, and men making somewhat higher offers for the latter, giving some credence to the idea that any such gender differences may be context-dependent.[4]

The paper most closely related to this paper is Brenner and Vriend (2006). They considered a one-player UG in which the experimental subjects play the role of proposers, while the receiving players are represented by a computer algorithm using fixed acceptance probabilities for all possible offers. This approach allowed to strip away some of the potential explanations for the behavior of players in ultimatum games. More specifically, as there are no other players, other-regarding preferences and fairness lose most of their relevance.[5]

This approach also eliminated the relative speed of learning issue as there is no learning by the receivers. In some sense, one can simply consider fixing the behavior of the receivers as the limit case of what is typically observed in ultimatum games (that proposers learn more quickly than receivers due to the incentives faced). But fixing the receivers' behavior gave some additional advantages. For a start, it gave full experimental control of the exact acceptance probabilities used by the receivers. This allowed them to determine exactly the behavior of receivers, e.g. to make sure that there is monotonicity of acceptance rates and that the monetary payoff maximizing offer is the minimum offer,

[4]As suggested also by Li et al. (2018).

[5]While trying to eliminate such factors (and related factors such as emotional ones or reputational concerns) as much as possible through the design, of course one cannot exclude that the behavior of some proposers may still be influenced by some residual concerns along these lines.

which also made the learning task of proposers as simple as possible. It also allowed some control of the beliefs of the proposers, as they were told that acceptance probabilities were fixed and monotonously increasing in offer sizes. This simplified the learning task for proposers, as they knew that they only had to learn about the fixed behavior of receivers.

In this paper we are building on this approach of Brenner and Vriend (2006). This also means, to get back to the gender issue in ultimatum games, that we will avoid two complications seen in the literature described above. First, as we will abstract from the role of the receivers, we can focus on the gender effect for proposers as such, without this being confounded with effects related to the precise gender configuration for proposers and receivers or the information that players may or may not have about this. Second, we will also focus on the usual (neutral) way in which decisions to be made in UG experiments tend to be described to the subjects, avoiding any kind of 'rich' framing.

2.2 Literature: Probability Matching (PM)

Probability Matching (PM) is a well-documented anomaly in the handling of probabilities. Imagine a player being asked to guess repeatedly the value of an independently drawn random number, which will be 1 or 0 with probabilities p and $(1 - p)$ respectively, with $p > 0.5$. Every correct answer, which means the player's guess is the same as the generated random number, leads to a fixed payoff while this will be 0 for a wrong answer. Payoff-maximizing behavior would imply choosing option 1 consistently. But this is not what is typically observed, as the allocation of trials tends to match the probabilities of success, with individuals choosing a fraction p of their decisions option 1 and a fraction $1 - p$ option 0.[6]

There are some significant differences between the choices of male and female players (see e.g. Gal and Baron, 1996; West and Stanovich, 2003). For example, Gal and Baron (1996) reported 69.9% of men choosing payoff-maximizing responses, i.e. not probability matching, whereas only 36.2% of women did so.

While the performance on various cognitive tests and their link to decision-making biases have been reported in the literature,[7] including some related to the

[6]See e.g. Grant et al. (1951), Goodnow (1955), Neimark and Shuford (1959), Gal and Baron (1996), Vulkan (2000), Rubinstein (2002), Shanks et al. (2002), West and Stanovich (2003), or Thaler (2016). This is true no matter whether these probabilities had been given upfront or whether they needed to be learned (see e.g. Koehler and James, 2010).

[7]See e.g. Oechssler et al. (2009), Georganas et al. (2015), Cueva et al. (2016), or Bosch-Rosa et al. (2018).

handling of probabilities, there has been relatively little on the relation between PM performance and behavior in games. An exception is Kim and Kim (2021), who link PM to behavior in the '$11-20$ Money Request Game' (Arad and Rubinstein, 2012).

For our context of analyzing UG behavior, a PM test seems particularly relevant because the underlying decision situation in a PM test and the UG is very similar. Both can be seen as sharing the underlying structure of a 'bandit problem'.[8] The PM test can be presented as the following two-armed bandit problem: For arm 1, a player will receive 1 with a chance of p and 0 with a chance of $1-p$, whereas for arm 0, a player will receive 1 with a chance of $1-p$ and 0 with a chance of p. Similarly, the UG can be presented as a multi-armed bandit problem: For each arm x, a player will receive $(1-x)$ with a chance of $p(x)$ and 0 with a chance of $1-p(x)$.

2.3 Literature: Risk attitudes (RA)

We use the 'off-the'-shelf' risk attitude test from Holt and Laury (2002). This allows estimating the value of a risk-aversion parameter of a utility function characterized by constant relative risk aversion, and thereby to distinguish e.g. risk-seeking from risk-neutral subjects. Charness et al. (2020) provide a discussion of a range of alternative ways to measure risk attitudes.

The relation between risk attitude and behavior in the UG has received relatively little attention in the literature.[9] In Eckel and Grossman (2001), the role of risk is confined to the background, when discussing gender differences in the UG. Garcia-Gallego et al. (2012) analyze behavior in the UG and separately also measure the subjects' risk attitude. They find a positive effect from risk-aversion to UG offers only in a treatment in which the UG is framed as salary negotiation between employer and employee and the task to be rewarded is a real one. Berninghaus et al. (2015) find that self-reporting risk preferences before playing an UG significantly decreases the proposers' offers in the UG,[10] and that more risk-averse subjects make higher offers. Candelo et al. (2019) also combine UG experiments with measuring risk attitudes, but find no significant risk effect.

As to the role of gender when it comes to risk attitudes, the picture

[8] We did not actually use the 'bandit' framing in our experimental design, neither for the PM test nor for the UG, thus avoiding possible framing and spillover effects.

[9] Interestingly, the word "risk" is not even mentioned in the Güth and Kocher (2014) survey. And the UG is not mentioned at all in Machina and Viscusi (2014).

[10] We always measured the RA after the UG.

found in the literature seems somewhat mixed, with many studies confirming the stereotype of women being more risk-averse, but also a good number showing no significant difference.[11] Surveying the literature, Holt and Laury (2014) conjecture that what may be key is "the application of expected value calculations" (p. 180), with on the one hand some individuals ("educated males in developed economies", p. 181) being more adept in this sense, and on the other hand some elicitation methods making it harder to calculate such expected values than other methods.

3 Experimental design

In our experiment each subject carried out three tasks: A one-player Ultimatum Game (UG), a Probability Matching (PM) test, and a Risk Attitude (RA) test. Details of these three tasks will be presented in the next few subsections. In addition, at the end of each session we had a questionnaire to collect some basic information, including sex, age, grade, and field of study. We used the oTree (Chen et al., 2016) platform for this experiment.

3.1 Design: Ultimatum Game (UG)

The UG task itself is based on Brenner and Vriend (2006), focusing again on the behavior of the proposers in ultimatum games. Whereas in Brenner and Vriend (2006) we focused on the learning behavior of proposers, eliminating concerns about fairness and other-regarding preferences as well as coevolutionary learning processes, here we also we remove the need for learning by the proposers. Thus, we continue our research strategy of stripping down standard ultimatum game experiments to eliminate potential explanations for the behavior of proposers.

To do this, in our main treatment (the Explicit treatment), subjects are told upfront the exact acceptance probabilities for each possible offer in the UG. Subjects in the Implicit treatment, on the other hand, need to learn these probabilities. That is, we use the information about the acceptance probabilities by the receivers as a treatment variable. Each subject participates in only one of these two treatments. Thus, while we adopt a "within-subjects" design for the three tasks, that is, each subject undergoes UG, PM as well as RA, we

[11] For the former see e.g. Eckel and Grossman (2002, 2008), Charness and Gneezy (2012), Garcia-Gallego et al. (2012) or Noussair et al. (2014), and for the latter e.g. Holt and Laury (2002), Andersen et al. (2006), Harrison et al. (2007), or Chakravarty et al. (2011).

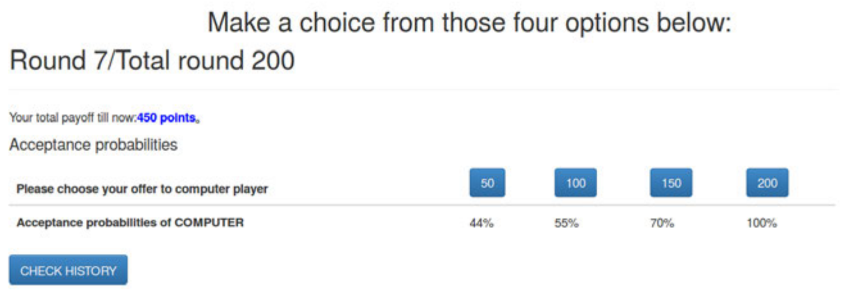


Figure 1: Decision making screen for UG in the Explicit treatment

use a “between-subjects” design with respect to the treatment variable whether information about acceptance probabilities is made explicit or kept implicit. The design of the UG task in the Implicit treatment in our experiment corresponds largely to the experiment reported by Brenner and Vriend (2006), and we can think of the Implicit treatment here as a control treatment.

In either treatment, there are 200 rounds in which the subjects play the role of proposers in an UG against receivers who are computer players. The amount available is 200, with the possible offers being 50, 100, 150, and 200. The fixed acceptance probabilities used by the receivers (unknown to the proposers in the Implicit treatment) are 0.44, 0.55, 0.70 and 1.00 for offers of 50, 100, 150 and 200, respectively, and the corresponding expected payoffs are 66, 55, 35, and 0.

Thus, in line with what one would expect in standard UG experiments, acceptance probabilities increase monotonically with offer size, whereas in line with a standard UG the expected monetary payoffs monotonically decrease with offer size. As a result, just as in the subgame perfect Nash equilibrium of the corresponding standard UG, the payoff maximizing offer would be to choose the minimum offer of 50.[12]

Giving the acceptance probabilities for the UG task in the Explicit treatment is the only difference with the Implicit treatment. These acceptance probabilities are given on the decision screen every time that a player has to make their choice from the available options in the Explicit treatment (see Figure 1). In either treatment subjects could see their history by clicking on the “Check history” button. Detailed experimental instructions can be found in Appendix A.

Before proceeding to the start of the experiment, after reading the instructions, each player was tested for their understanding of the UG setup,

[12]Compared to Brenner and Vriend (2006) we have fewer options and larger differences between the options (the expected payoff difference between two options is at least 16.6%), making the task easier.

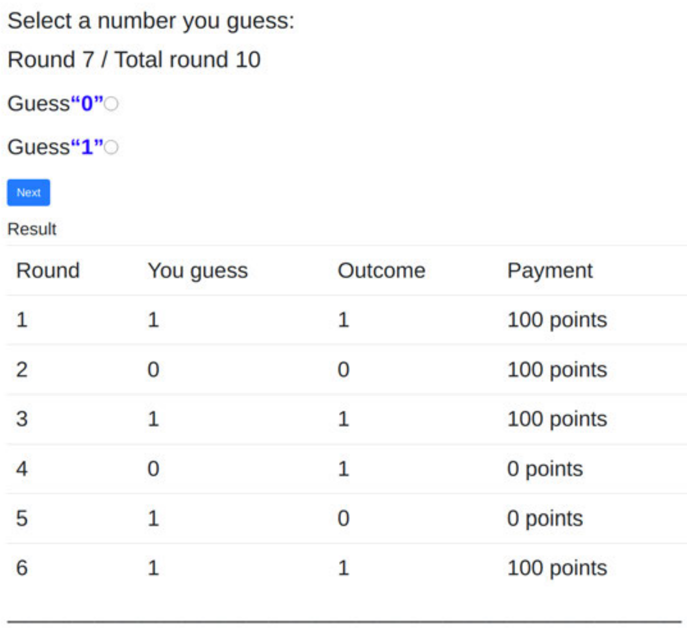


Figure 2: Decision making screen for PM test

including the fact that their opponent was a computer, and the available payoffs for proposers and receivers. We also reminded the subjects that as a computer player their opponent cannot receive real payments, and that the acceptance probabilities were fixed and would not adjust depending on the actions of the human players.

3.2 Design: Probability Matching (PM)

In the Probability Matching (PM) task, subjects are told that there is a random number that will be 1 with probability $p = 0.7$ and 0 with probability $1-p = 0.3$. A subject needs to guess the value of this random number. When it happens to be the same as the generated number, the player receives 100 points. Each subject did the PM test for 10 rounds, with the random values being i.i.d. Note that for any risk attitude, the payoff maximizing choice is to choose option 1 for each of the 10 rounds. Figure 2 shows the decision-making screen for the PM test. Subjects could check their entire history of choices and outcomes directly in the lower part of the screen. There is no difference between the PM task in the Implicit and Explicit treatment.

Note that in both PM test and UG the subjects need to make some repeated decisions, with the probabilities of success fixed and known to be fixed. In the

PM test and the Explicit treatment of the UG these probabilities are also given upfront to the subjects. A minor difference with our UG is that in the PM test the payoffs for success are identical for both options, whereas in the UG they depend on the exact offer made, which means that the PM test is simpler. Therefore if dealing with probabilities in the PM test is problematic for some subject, then this is most likely going to be problematic in the UG as well.

3.3 Design: Risk attitudes (RA)

To measure the subjects' risk attitude we used a standard Holt and Laury (2002) test. Subjects were asked to make choices for ten situations. Each situation consisted of a choice between paired lotteries, a relatively safe lottery A and a riskier lottery B. Figure 3 shows the screen where the subjects enter their decisions. In each situation, the probability to obtain the high (low) price is the same for lotteries A and B, and the essential difference is that the high price is higher and the low price is lower in the riskier lottery B. The high and low prices themselves are constant across the ten decision situations for both lotteries A and B.

The probability to obtain the high price increases from 1/10 in the first situation to 10/10 in the tenth situation. Thus, as we go through the range of ten situations, the riskier lottery B becomes more attractive as the probability of the high price increases and this high price is higher in lottery B than in lottery A. Eventually, when we reach the final of the ten choice situations, lottery B gives a higher monetary payoff than lottery A with probability 1, which means that everybody will prefer lottery B no matter how risk-averse they may be. For the earlier nine choice situations, the optimal choice between the safe lottery A and the riskier lottery B depends on the individual's risk attitude.

For the first choice situation, only extreme risk-seekers will go for lottery B. But as we go through the range of ten situations, and lottery B becomes more attractive, each individual's preference will at some point switch from the safe lottery A to the risky lottery B. Risk-neutral individuals will switch from the safe lottery A to the risky lottery B right after situation 4, as the expected monetary payoffs of lotteries A are greater than those of lotteries B up to that point while the opposite is true for situations 5 to 10. Risk-seeking individuals will switch (weakly) earlier and risk-averse individuals will switch (weakly) later than risk-neutral ones. Thus, the point where individuals make this switch gives us some information about their risk attitude. Holt and Laury (2002) use these switching points to compute for each subject the value of a risk-aversion parameter of a utility function characterized by constant relative risk aversion.

For the RA test one of the ten choices would be selected at random for a player to determine the monetary reward after the experiment, with the monetary amounts involved being the figures listed on the screen. Before making their choices we tested their understanding of the procedure determining the monetary rewards. Just like with the PM test, also for the RA test there was no difference between the Explicit and Implicit treatment.

3.4 Design: Summary descriptive statistics experimental subject pool and design

Table 1 summarizes some descriptive statistics concerning our experimental subject pool and design.[13] In total, we have 634 participants, with each of them participating only in one session, and hence in only one treatment. The larger number of subjects in the Explicit treatment reflects the idea that the Implicit treatment could be seen as a control treatment. For the first two tasks, some sessions used the UG-PM order, whereas other sessions used the PM-UG order, with about half of the subjects for each of these orders. The RA test was always done last. The subject pool consisted of students from Beijing Normal University (BNU), Zhejiang University (ZJU), and Zhejiang Gongshang University (ZJSU), in proportions of approximately 3/6, 2/6 and 1/6 respectively, with the grades most of them were studying for fairly evenly spread over the various years of undergraduate and Master's programmes. The age distribution reflects this as well, with the average age being 21.7 years and the majority (97.6%) being between 18 and 26. Almost 1/5 of them were studying on some Economics, Business or Finance programme (broadly defined). Most (55%) of them had not participated in any experiments before, some (33%) in one, and only a few in two (10%) or three (1%). Almost half (47.6%) of the participants were male and just over half (52.4%) female students.

As explained above, the players were rewarded on the basis of the experimental points accumulated in the UG and PM test, with 200 points corresponding to 1.0 CNY. In addition, they were paid according to the realization of one (randomly selected) of the ten lottery choices of the RA test. The average duration of the sessions was just over one hour, and the average pay

[13] While all variables will be described as they appear in the main text, a complete and detailed list of the variable definitions can be found in Appendix B.

Make a decision for every situation

	A	B	Your Choice
Situation 1	1/10 of ¥2.00 9/10 of ¥1.60	1/10 of ¥3.85 9/10 of ¥0.10	Decision 1: ○ A ○ B
Situation 2	2/10 of ¥2.00 8/10 of ¥1.60	2/10 of ¥3.85 8/10 of ¥0.10	Decision 2: ○ A ○ B
Situation 3	3/10 of ¥2.00 7/10 of ¥1.60	3/10 of ¥3.85 7/10 of ¥0.10	Decision 3: ○ A ○ B
Situation 4	4/10 of ¥2.00 6/10 of ¥1.60	4/10 of ¥3.85 6/10 of ¥0.10	Decision 4: ○ A ○ B
Situation 5	5/10 of ¥2.00 5/10 of ¥1.60	5/10 of ¥3.85 5/10 of ¥0.10	Decision 5: ○ A ○ B
Situation 6	6/10 of ¥2.00 4/10 of ¥1.60	6/10 of ¥3.85 4/10 of ¥0.10	Decision 6: ○ A ○ B
Situation 7	7/10 of ¥2.00 3/10 of ¥1.60	7/10 of ¥3.85 3/10 of ¥0.10	Decision 7: ○ A ○ B
Situation 8	8/10 of ¥2.00 2/10 of ¥1.60	8/10 of ¥3.85 2/10 of ¥0.10	Decision 8: ○ A ○ B
Situation 9	9/10 of ¥2.00 1/10 of ¥1.60	9/10 of ¥3.85 1/10 of ¥0.10	Decision 9: ○ A ○ B
Situation 10	10/10 of ¥2.00 0/10 of ¥1.60	10/10 of ¥3.85 0/10 of ¥0.10	Decision 10: ○ A ○ B

Figure 3: Paired lotteries for RA test

Variable	**Category**	#
Subjects	All	634
Treatment	Implicit	209
	Explicit	425
Order tasks	UG-PM-RA	311
	PM-UG-RA	323
University	BNU	304
	ZJU	217
	ZJSU	113
Grade	0 Freshman	109
	1 Sophomore	101
	2 Junior	82
	3 Senior	81
	4-6 Master's	239
	7-9 PhD	17
	10 Other	5
Degree	Economics/Finance/Business	113
	other	521
Gender	Male	302
	Female	332
Age	-17	3
	18	41
	19	75
	20	108
	21	87
	22	84
	23	80
	24	76
	25	48
	26	20
	27+	12
	(avg.	21.7)
Participations	0	351
	1	212
	2	63
	3	8
	(avg.	0.57)

Table 1: Summary descriptive statistics experimental subject pool and design

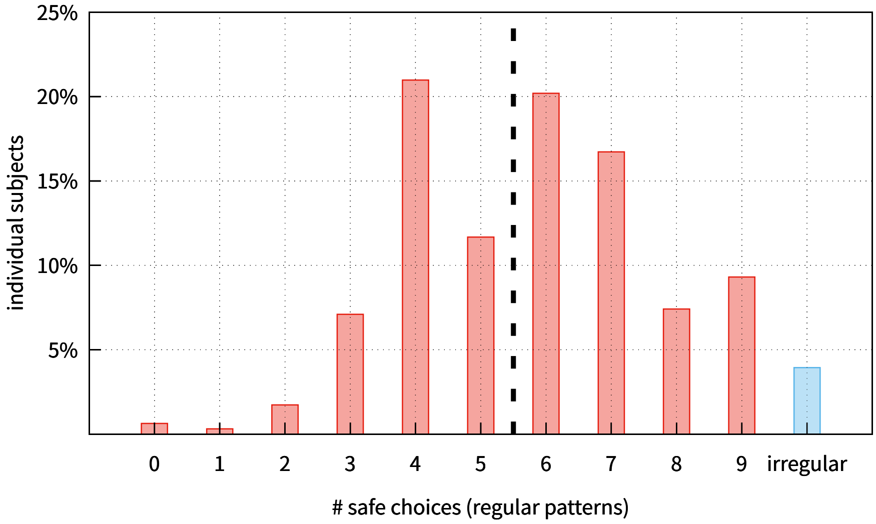


Figure 4: Frequency distribution of the # of safe choices in the RA test

was about 68 CNY.

4 Analysis

In Subsection 4.1 we will provide a first overview of the data, focusing on the main variables concerning the behavior of the subjects in the experiment, while Subsection 4.2 will present some detailed statistical (regression) analysis.

4.1 Analysis: Overview

Before we consider the behavior of the subjects in the UG itself, we first have a look at the behavior in the RA and PM tests. We begin with the RA test. Figure 4 shows the frequency distribution of the number of safe choices made by the individual subjects. As explained in Section 3, the choice pattern to be expected here is a regular one, in which subjects may start with the safe lottery A or the riskier lottery B, and as we move through the range of choice situations and the lotteries B become more attractive, for each subject there will come a point where they prefer to switch to the riskier lotteries B and then stick with those, with the exact switching point depending on their risk attitude.

Just like in Holt and Laury (2002), we find that there are some players who display an irregular choice pattern in the RA test. Such irregular patterns may be characterized by multiple switching points, by switching in the wrong direction (from risky to safe lotteries), or by failing to choose the 'risky' lottery even when the higher payoff is realized with certainty.[14] We deal with any irregular choice patterns as follows. If flipping one single safe (risky) choice would lead to a regular pattern with a unique switching point, then we carry out this flip and re-count the corresponding number of safe choices.[15]

As we can see in Figure 4, the mode (with 21%) consists of subjects making 4 safe choices before switching to the riskier lotteries, which corresponds to risk-neutrality. With fewer than 4 safe choices, we find the risk-seeking subjects (10%), whereas we find 12% mildly risk-averse subjects (with 5 safe choices), and 54% of the subjects are more seriously risk-averse (more than 5 safe choices).

As mentioned in Section 3, the switching points of the Holt and Laury (2002) RA test can be used to compute the corresponding risk-aversion parameter of a CRRA utility function, which can, then, be used to work out what offer would be optimal for the individuals concerned in the UG. Recognizing that alternative functional specifications and alternative tests may be relevant, we create a dummy variable *RA pass* to simply differentiate between those for whom the RA test provides evidence that making the minimum UG offer of 50 would be optimal, and those for whom we do not have such evidence.

The former include risk-seeking, risk-neutral and mildly risk-averse subjects with a regular RA test performance, i.e. players with 0 up to and including 5 safe lottery choices before switching to the risky lotteries. We consider those subjects to have passed the RA test, and hence we will use in our analysis $RA\ pass = 1$ for them (the 42% to the left of the dashed line in Figure 4). The other category of subjects includes the more seriously risk-averse individuals, i.e., those with a later switching point and those with seriously irregular choice patterns, for whom we will use in our analysis instead $RA\ pass = 0$ (the 58% to the right of the dashed line).

Next, the PM test. Figure 5 shows the frequency distribution of the number of optimal choices made in the PM test by the individual subjects. As we can see, almost half (48%) of the subjects consistently made the optimal choice, whereas the other subjects went for some mix of the two options, on average making the

[14] Holt and Laury (2002) ignore all such irregularities and simply count the number of safe choices (whatever the exact pattern), arguing that in many cases irregular patterns seem merely the result of some hesitation around a unique switching point.

[15] Note that in some cases this will imply flipping a single safe as well as a single risky choice at the same time. Note also that this procedures leads to players with 10 safe choices being classified like those with 9 safe choices (having flipped their final choice). Before this correction in the RA choice data there were 89 irregular patterns (14%), and after correction 25 (4%).

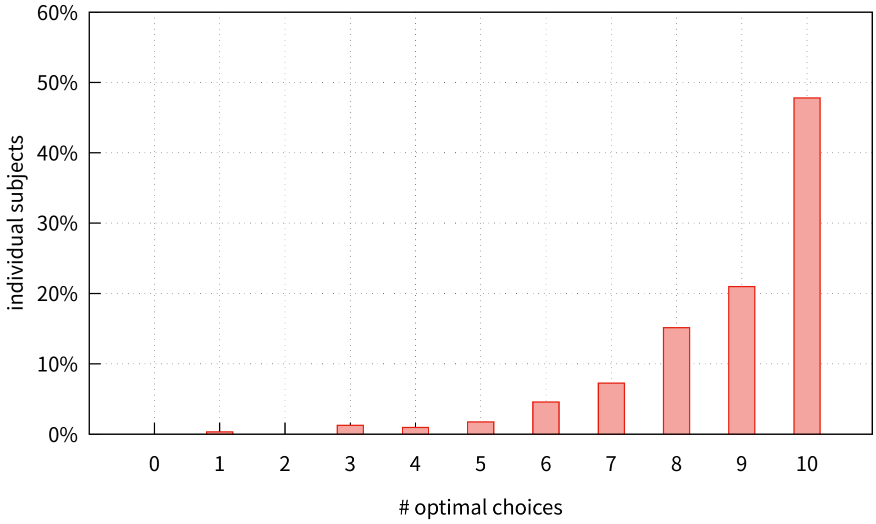


Figure 5: Frequency distribution of the # of optimal choices in the PM test

optimal choice 78% of the time and the non-optimal choice 22%.[16]

To characterize the difference between subjects making consistently optimal choices and those showing poor processing and handling of probabilities in the PM test, we define a dummy variable *PM pass*. This dummy is equal to 1 for players making consistently optimal choices, thereby passing the Probability Matching (PM) test, and equal to 0 for the other players.

We now turn our attention to the behavior in the UG itself. The bars in Figure 6 show the average offers in the Implicit and Explicit treatments of the UG (for each treatment averaged over all players and all periods). The inset shows how these average offers evolve over the 200 periods.

We make the following initial observations. First, the average offers in the Explicit treatment are lower than in the Implicit treatment (the average over all periods is 74.4 against 90.8). Second, the average offers slowly decrease over time in a very similar way in both treatments. For example, in the Implicit treatment the average over periods 1-50 is 92.6, whereas it is 88.5 over periods 151-200, and in the Explicit treatment the average decreases from 76.8 to 72.3 in the same periods.[17] Third, the average offers in both treatments remain substantially

[16]Given the probabilities of success for the two options, perfect Probability Matching would have implied an average of 70% optimal choices and 30% non-optimal choices.

[17]We are agnostic as to whether these changes over time are due to a learning effect or an endowment effect. In the Implicit treatment it may appear more 'natural' to think of this as a learning effect, but in either treatment we cannot rule out either or even a combination of

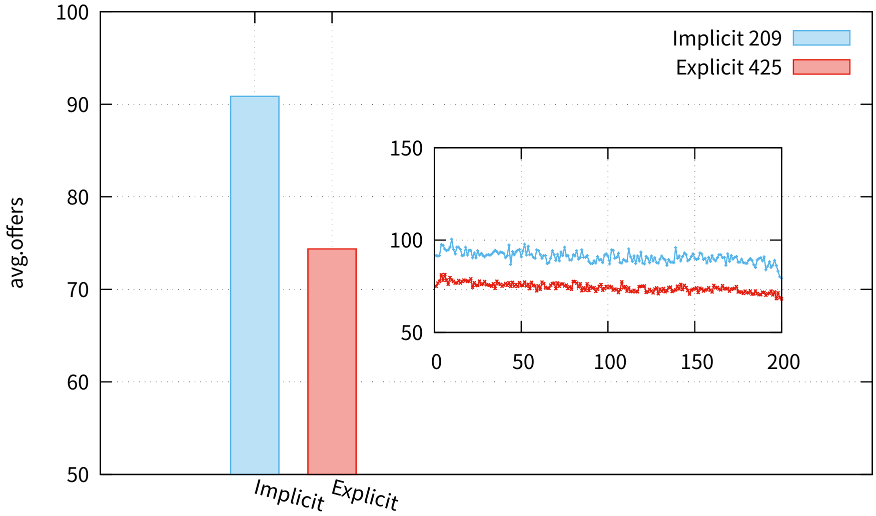


Figure 6: UG offers in Implicit and Explicit treatments

above the minimum offer of 50. In the Implicit treatment the average offer made over periods 151-200 is about 44% of the available pie size of 200,[18] whereas even in the Explicit treatment it still is about 36% of the available pie size.

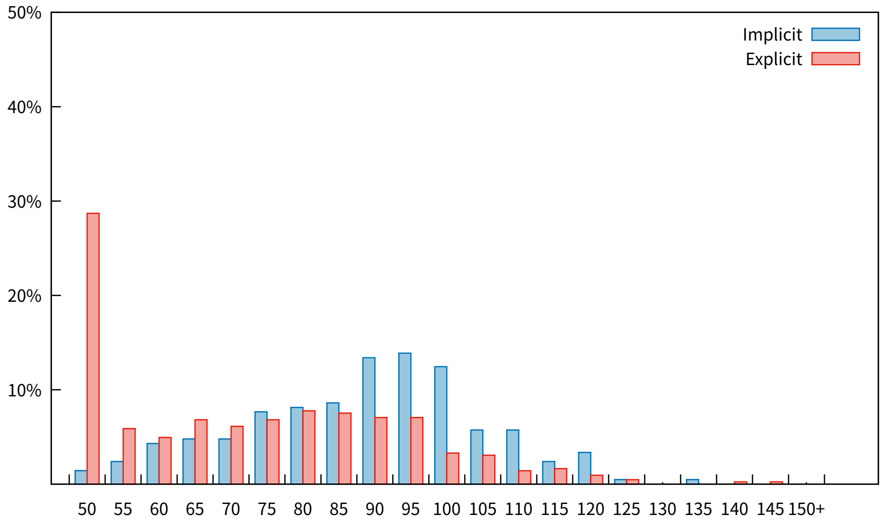
Thus, in an experimental design aimed at abstracting from other-regarding preferences and (coevolutionary) learning processes, players very slowly move their offers towards the minimum offer, but even when the acceptance probabilities are explicitly given, the average offers are still substantially above the payoff maximizing minimum offer of 50 after 200 periods.

These results confirm that the learning about acceptance probabilities in ultimatum games may be important, and also that it may be very slow, such that average offers should not be expected to be near the payoff maximizing minimum offer for a very long time. But these results also show that it might be that this is not the sole reason for the observed non-maximizing behavior, since even when the acceptance probabilities $p(x)$ are explicitly given, and hence the players no longer need to learn them, they still demonstrate non-maximizing behavior.

There is, however, some important heterogeneity underlying these slowly evolving average offers. Taking for each individual player their average offer over the 200 periods, Figure 7 shows the distribution of offers in the Implicit and Explicit treatment of the UG. We see that there is some similarity in the

both effects.

[18]This is essentially in line with the findings of Brenner and Vriend (2006).



Note: bin “50” = [50, 55), “55” = [55, 60), . . . , . . . , “150+” = [150, 200]

Figure 7: Distribution of UG offers in Implicit and Explicit treatments

distributions of the two treatments. By far most of the average offers are in the 50-100 range, with only some in the 100-150 range, and none above 150. If we ignore for a moment the first bin labeled “50”, the distributions also look similar in that the mode appears to be just below 100, or 50% of the pie size. In some sense, one could argue that the distribution of the Explicit treatment appears simply a somewhat shifted-to-the-left version of that for the Implicit treatment. But there is one important additional feature to note. In the Explicit treatment there is a substantial number of players (29%) who have an average offer strictly below 55 (the first bin), and most of those (60%) consistently chose an offer of 50 throughout all the 200 periods. In the Implicit treatment only 1% of the players had an average offer corresponding to the first bin (below 55), and not one single player consistently chose an offer of 50.

Therefore, we need to investigate this heterogeneity. More in particular, we want to see how the average UG offers that individual players make may be related to their attitude with respect to risk, as well as their ability to interpret and deal with probabilities, and the differences therein. Table 2 shows some first evidence linking the players’ average UG offers to their RA and PM test performance, and in addition the table also distinguishes (for reasons that will become apparent) the players’ gender. For each combination of these variables of interest, and distinguishing the two treatments, the table shows the number of

Treatment		Implicit			Explicit		
		n	avg	sd	n	avg	sd
	all	209	90.8	16.7	425	74.4	20.8
RA test	pass	94	89.2	16.0	175	67.4	18.7
	fail	115	92.2	17.2	250	79.2	20.9
PM test	pass	99	89.2	18.3	204	66.9	19.6
	fail	110	92.3	15.1	221	81.3	19.5
Gender	male	96	89.2	17.1	206	68.7	19.8
	female	113	92.2	16.3	219	79.6	20.4

Note: underlying variable is UG offers of individual players (averaged over periods 1-200)

Table 2: UG offers by treatment and category

players, the average UG offer by these players, as well as the standard deviation.[19]

We can make some first observations. First, the RA and PM tests and the *Gender* variable each split the subject pool in approximately equal groups, with 42%, 48% and 48% of the players passing the RA test, passing the RA test or being male, respectively. Second, that UG offers tend to be higher in the Implicit treatment than in the Explicit treatment is confirmed for all groups distinguished in the table. Third, in both treatments, players failing the RA test or the PM test, as well as female players offer more in the UG.

For some further insight concerning these variables see Appendix C, where we present a number of graphs analogous to Figures 6 and 7, showing with respect to each of these variables how the average offers evolve over time as well as the frequency distribution of these average offers. The main observation there is that while all these average offers evolve in a similar way to what we saw in Figure 6, and also the distributions of the average offers with respect to the variables concerned look similar to what we saw in Figure 7 for the Implicit and Explicit treatments, there is one outstanding feature. And that is that in the Explicit treatment among the players passing the RA test, those passing the PM test as well as male players there is a much higher frequency (about 40%) of average offers corresponding to the [50, 55) bin than for those failing these tests or female players (about 20%).

The players' performance in the RA test and PM test is not statistically

[19]See Appendix B for a description of all variables.

independent.[20] This seems unsurprising as both tests relate to how players deal with probabilities. Notwithstanding this statistical dependence, the correlation between these two tests is relatively limited. With perfect (positive) correlation the number of mixed profiles would have been 0 (with each player either passing both tests or failing both), whereas if the RA and PM tests had been independent the expected number of subjects with a mixed test profile (passing one of the two tests while failing the other) would have been 315 (out of 634). We found instead 274 players with mixed test results, much closer to the independent than to the perfect correlation benchmark, which suggests that the RA and PM tests capture somewhat different aspects of dealing with probabilities.

RA test performance and *Gender*, as well as PM test performance and *Gender* are not independent either.[21] Whereas 35% of female players pass the RA test, this is 50% for male players, and while 40% of female players pass the PM test this is the case for 56% of male players.

Thus, we find that female players offer more in the UG game, and that female player are more risk-averse and tend to fail the PM test more often. A question, then, is whether different risk attitudes and differences in the PM test performance capture all the differences between male and female players as far as UG offers are concerned. This will be one of the questions considered in the next subsection, where we present our regression analysis.

4.2 Regression Analysis

In this subsection we present a regression analysis of the behavior of the players in the UG. The main variables of interest are those directly related to how the players deal with stochasticity, i.e., the players' RA and PM test performance. Besides these focus variables, we take into account a couple of secondary design variables as well as some individual (demographic) characteristics obtained from a post-experiment questionnaire. We estimate the players' predisposition to offer to the other player in the UG as a linear function of their performance in the Risk Attitude test (RA) and in the Probability Matching (PM) test, the order of their tasks, their university, their field of study, current grade (level), as well as their age and gender, and their participations in previous experiments.

Before examining the results of our econometric analysis, we first briefly discuss the dependent as well as the independent variables as used in our

[20] χ^2-test, significant at 1%.

[21] Both χ^2-test, significant at 1%.

regression analysis, and we present our hypotheses concerning the links between these independent and dependent variables.[22]

The dependent variable we focus on is the individual UG offers, for each player averaged over their entire history of 200 periods.[23] As average offers are bound by the minimum and maximum possible offers in the UG (50 and 200, respectively), we use the Tobit regression model.[24] In our regression analysis we consider nine independent variables. We start with the two focus variables.

RA pass: As explained in Subsection 4.1, this is a dummy variable, equal to 1 if the RA test indicates that for this player the optimal UG offer is equal to the minimum offer of 50, and equal to 0 if we have no such evidence. Our hypothesis was that players for whom the optimal UG offer corresponds to the minimum offer of 50 would offer less than other players.

PM pass: This is another dummy variable, equal to 1 for players making exclusively optimal choices, thereby passing the Probability Matching (PM) test, and equal to 0 for all other players (see Subsection 4.1). Our hypothesis was that poor processing and handling of probabilities was one of the causes of UG offers above the minimum offer, and thus we expected players passing the PM test to make lower UG offers.

We next consider two secondary design variables (see also Section 3).

Order tasks: This is a dummy variable, equal to 1 if the order was PM-UG-RA, and equal to 0 if the order was UG-PM-RA. Our hypothesis was that starting with the PM might help some players to choose lower UG offers, as they would deal more carefully with probabilities in the UG as a result of some priming effect.

University: With Explicit treatment sessions conducted at BNU and ZJU, and Implicit sessions at BNU and ZJSU, this dummy variable was equal to 0 for BNU players (in either treatment), and 1 for ZJU or ZJSU players (in Explicit and Implicit treatments, respectively). We had no particular hypotheses concerning the differences between these universities.

Finally, we consider some demographic characteristics collected from the individuals in a post-experiment questionnaire.

Grade: This is a variable indicating the grade at which the players were

[22] See Appendix B for a complete list of all variables.

[23] In Tables 4 and 5 in Appendix D we show the robustness to alternative specifications of the dependent variable, reporting regression results where the dependent variable is computed instead over periods 1-10, 1-50, 1-100, 101-200, 151-200 and 191-200, respectively.

[24] As it is, only in the Explicit treatment there was some left-censoring (73 out of 425 observations), and no right-censoring occurred in either treatment.

currently studying, ranging from 0 ('Freshman') to 9 ('PhD Third grade'), and 10 ('Other').[25] Our hypothesis was that players studying at a higher grade would be more sophisticated and hence make lower UG offers.

Age difference: As the players' age was (unsurprisingly) highly correlated with their grade, we constructed the variable *Age difference* by taking for each player the difference between their age and the average age of players at their grade. Our hypothesis was that players who are older than the average of their peers would be less sophisticated and hence make higher UG offers.

Participations: This variable indicates the number of previous participations in experiments. Our hypothesis was that more experienced players would make lower UG offers.

Degree: This dummy variable indicates whether the player was studying for a degree in the field of Economics, Finance or Business Studies in a broad sense (value of 1), or not (value of 0). Our hypothesis was that students in the field of Economics, Finance or Business Studies would make lower UG offers.

Gender: A dummy variable equal to 1 (male) or 0 (female). Our hypothesis was that female players would make higher UG offers, based on evidence from the literature on bargaining games, where female players tend to make higher offers. Experimental evidence also suggests that female players tend to be more risk-averse, and tend to perform worse in a PM test. Controlling for RA and PM test performance, our hypothesis was that there would be no residual gender effect as far as UG offers are concerned.

Table 3 shows the Tobit regression results for the Implicit treatment (left-hand side of the table) and the Explicit treatment (right-hand side). We present three specifications for each treatment, with all estimated coefficients, and the standard errors in brackets. For the Implicit treatment, column (1) shows the regression including only the two focus variables *RA pass* and *PM pass* as independent variables in addition to a constant. Column (2) is the complete specification, including also all control variables, i.e. the secondary design as well as the questionnaire variables. Column (3) is a more parsimonious specification, which results after elimination of insignificant variables.[26] Columns (4) to (6) are the corresponding specifications for the Explicit treatment.

We start with the Explicit treatment (right-hand side of the table). With

[25]See Appendix B for the complete list.

[26]For columns (1) and (3), the F-statistics and corresponding p-values for the joint restrictions with respect to the complete specification in column (2) are reported in the table.

	Dependent variable: individual UG offers averaged over periods 1-200					
	(1)	(2)	(3)	(4)	(5)	(6)
Treatment	Implicit	Implicit	Implicit	Explicit	Explicit	Explicit
Intercept	93.5783***	96.2184***	97.7644***	84.4023***	91.8264***	89.7934***
	(1.8593)	*(3.3687)*	*(2.4096)*	*(1.6590)*	*(3.0246)*	*(2.1114)*
RA pass	-2.8854	-3.7751	−4.2741*	−11.3110***	−10.5567***	−9.8625***
	(2.3032)	*(2.3073)*	*(2.2906)*	*(2.2493)*	*(2.2290)*	*(2.2103)*
PM pass	-3.0241	-1.9957		−15.9920***	−14.7548***	−15.1167***
	(2.2947)	*(2.3230)*		*(2.2144)*	*(2.1790)*	*(2.1842)*
Order tasks		2.8680			−3.9012*	−3.7600*
		(2.3134)			*(2.1391)*	*(2.1336)*
University		-0.7944			-2.2746	
		(2.7455)			*(2.2379)*	
Grade (level)		−1.3866**	−1.5644***		0.1728	
		(0.6137)	*(0.5958)*		*(0.4940)*	
Age difference		2.0222	2.0461*		1.5359*	1.4884*
		(1.2381)	*(1.2086)*		*(0.8856)*	*(0.8850)*
Participations		1.4107			-2.1746	
		(1.5681)			*(1.5823)*	
Degree (field)		1.2407			-0.8261	
		(2.6960)			*(3.9550)*	
Gender		-1.6179			−9.1959***	−9.1978***
		(2.3165)			*(2.1889)*	*(2.1751)*
Observations	209	209	209	425	425	425
LR χ^2	3.44	15.57	11.53	84.66	112.08	107.71
(degr.freed.)	*(2)*	*(9)*	*(3)*	*(2)*	*(9)*	*(5)*
(p-value)	*(0.1788)*	*(0.0765)*	*(0.0092)*	*(0.0000)*	*(0.0000)*	*(0.0000)*
F for restrictions	1.78	n.a.	0.68	4.03	n.a.	1.10
(degr.freed.)	*(7, 200)*	*n.a.*	*(6, 200)*	*(7, 416)*	*n.a.*	*(4, 416)*
(p-value)	*(0.0924)*	*n.a.*	*(0.6665)*	*(0.0003)*	*n.a.*	*(0.3572)*

Notes: All variables are described in the main text and in Appendix B. For each variable we report the estimated coefficient with the standard error in brackets. For each column, the number of observations corresponds to the number of individual players. For each column we report the Likelihood Ratio (LR) χ^2-test that at least one of the predictors' regression coefficients is not equal to zero (with the degrees of freedom and p-value in brackets). In columns (1), (3), (4) and (6) we also report the F-test for the joint restrictions, starting from the complete specifications in columns (2) and (5), with the corresponding degrees of freedom and p-value in brackets. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$

Table 3: Tobit regression results

an intercept of 84.4, column (4) shows that players who are not too risk-averse make lower UG offers (-11.3, significant at 1% level), and those who pass the PM test offer significantly less as well (-16.0, at 1%). Both are in line with our hypotheses.

The F-test for the joint restrictions in column (4) with respect to the complete specification in column (5) shows that the control variables cannot be completely ignored. Column (5) gives an insight as to which of these control variables matter. The main additional effect, interestingly, is that controlling for RA and PM performance and the gender differences therein, male players offer significantly less than female players in the UG (-9.2, at 1%). In addition, there is a weaker effect from the order of tasks, with those primed by the PM test making somewhat lower UG offers (-3.9, at 10%), and players who are older than the average for the grade at which they study offer slightly more (1.5/year, 10% significance). The grade (level) they are studying at, their field of study and the university do not appear to be significant, and neither do previous participations in experiments seem to matter. While these findings for the order effect and age difference are in line with our hypotheses, the strong gender effect is something that we had not anticipated. Notwithstanding these additional effects of the control variables, the complete specification in column (5) largely confirms what we find with the focus variables in column (4), with the coefficients and significance of the focus variables and the significance of the regression as such displaying similar values. Column (6) simply presents a more parsimonious regression obtained by eliminating as many variables as the F-test for joint restrictions would allow us. As we can see, coefficients and significances are not influenced too much.

The picture for the Implicit treatment (left-hand side of the table) is qualitatively similar, although some of the effects appear weaker than in the Explicit treatment. While the intercept in column (1) is somewhat larger (93.6), the focus variables *RA pass* and *PM pass* have reduced coefficients and lose their significances, but the signs are the same as for the Explicit treatment, with lower UG offers from players for whom the optimal UG offer was 50 as well as from players who passed the PM test. The regression as such does not appear significant (p-value of 0.18). Adding the control variables in column (2) does not change the picture much. Only the grade at which the players study is significant, with more advanced students offering slightly less in the UG (-1.4 per grade level, at 5%), while the regression as such becomes weakly significant (p-value of 0.08). Moving to the more parsimonious specification in column (3) gives a somewhat clearer picture, which is largely consistent with what we saw for the Explicit treatment. The effect of the RA test works in the same direction as in the Explicit treatment, with those for whom the optimal UG offer was 50 making indeed somewhat lower UG offers than the others (-4.3 at 10% significance). The PM test and *Gender* do not appear significant here, unlike in the Explicit treatment, although the sign is the same. The other minor effects or lack thereof seen in the Explicit treatment are largely confirmed here. There is again a minor effect with players older than the average at their grade level making higher UG offers (2.0/year at 10% significance), and there is a clear effect of grade level (-1.6/level at 1% significance). The weak effect of the order of tasks observed in the Explicit treatment disappears here, and previous experimental participations, field of study and university, again, do not matter.[27]

The key difference between the Implicit and Explicit treatment is that in the former the players do not know the underlying UG acceptance probabilities for the various offers. As a result, the level of uncertainty is much higher in the Implicit treatment, imposing a considerable learning task, and confusing players more, possibly overwhelming any other effects.[28] On average the individual players do indeed show a variance in their UG offers that is higher in the Implicit than in the Explicit treatment (999.7 vs. 554.1), and the size of the average period-to-period changes made in their offers is larger in the Implicit treatment as well (27.1 vs. 19.3). Besides possible confusion, these differences may also to some extent be a sign of the exploration and learning required in the Implicit treatment. Therefore, we consider another sign of possible confusion in the Implicit treatment.

For some players it is the case that while the RA test indicates that their optimal UG offer is 50, they do not choose this that often. If they choose that offer at most 19 times (strictly less than 10%) of the 200 periods we classify them as 'confused'. This is the case for 11 players in the Implicit treatment, whereas it happens with only 2 players in the Explicit treatment. If we exclude these 'confused' players, we get the regression results reported in Table 6 in Appendix E, showing exactly the same specifications as in Table 3. For the

[27]The p-value of the F-statistic to test all the joint restrictions when moving from column (2) to column (3) was 0.67.

[28]Note that these findings about the effects of the RA and PM test performance as well as gender in the Implicit and Explicit treatments seem in line with the conjectures by Holt and Laury (2014) about the relevance of expected value calculations in their discussion of the link between gender and risk attitudes.

Explicit treatment this merely confirms the results reported above. For the Implicit treatment, we see that both the RA and PM test become significant (albeit still with somewhat reduced coefficients and significances compared to the Explicit treatment), with players who are not too risk-averse and those passing the PM test offering again less in the UG.

5 Concluding remarks

We considered an UG game in which subjects play the role of proposers against computer opponents. Our main finding is that most players choose higher offers than the expected monetary payoff maximizing UG offer of 50 even when all acceptance probabilities are given upfront and even after 200 periods. We showed that this is significantly related to how they deal with probabilities, and all-in-all it seems that our hypotheses are largely confirmed. More precisely, this concerns two main effects. First, a substantial number of players offer more in the UG as they appear to be strongly risk-averse (RA test). Second, a similarly substantial number of players make higher than payoff maximizing offers as they appear to have some difficulty in dealing with probabilities (PM test). We also find some relatively weak minor effects related to some of the control variables (e.g. the order of tasks, grade, or age differences) that are in line with our hypotheses. In addition, we find a clear and strong UG effect of gender especially in the Explicit treatment. While controlling for their performance in RA and PM tests, we find that male players offer significantly less than female players, which requires further study.

References

Andersen, S., Harrison, G.W., Lau, M.I., Rutström, E.E., 2006. Elicitation using multiple price list formats. Experimental Economics 9, 383 405.

Arad, A., Rubinstein, A., 2012. The 11-20 money request game: A level-k reasoning study. American Economic Review 102, 3561 3573.

Berninghaus, S., Todorova, L., Vogt, B., 2015. A simple questionnaire can change everything: Are strategy choices in the coordination and ultimatum games stable?, in: 2015 IEEE International Conference on Smart

City/SocialCom/SustainCom together with DataCom 2015 and SC2 2015, pp. 452 456.

Bolton, G., Zwick, R., 1995. Anonymity versus punishment in ultimatum bargaining. Games and Economic Behavior 10, 95 121.

Bosch-Rosa, C., Meissner, T., Bosch-Domènech, A., 2018. Cognitive bubbles. Experimental Economics 21, 132 153.

Brenner, T., Vriend, N.J., 2006. On the behavior of proposers in ultimatum games. Journal of Economic Behavior & Organization 61, 617 631.

Camerer, C., 2003. Behavioral game theory: Experiments in strategic interaction. Princeton University Press.

Candelo, N., Eckel, C., Johnson, C., 2019. The proposer's behavior in the ultimatum game in 11 Mexican villages. Economics Letters 177, 5 8.

Carpenter, J.P., Burks, S., Verhoogen, E., 2005. Comparing students to workers: The effects of social framing on behavior in distribution games, in: Carpenter, J., Harrison, G., List, J. (Eds.), Field Experiments in Economics, Research in Experimental Economics, Volume 10. JAI/Elsevier, Amsterdam, Boston, pp. 261 290.

Chakravarty, S., Harrison, G.W., Haruvy, E.E., Rutström, E.E., 2011. Are you risk averse over other people's money? Southern Economic Journal 77, 901 913.

Charness, G., Garcia, T., Offerman, T., Villeval, M.C., 2020. Do measures of risk attitude in the laboratory predict behavior under risk in and outside of the laboratory? Journal of Risk and Uncertainty 60, 99 123.

Charness, G., Gneezy, U., 2012. Strong evidence for gender differences in risk taking. Journal of Economic Behavior & Organization 83, 50 58.

Chen, D.L., Schonger, M., Wickens, C., 2016. oTree - an open-source platform for laboratory, online, and field experiments. Journal of Behavioral and Experimental Finance 9, 88 97.

Chuah, S., Hoffmann, R., Jones, M., Williams, G., 2007. Do cultures clash? Evidence from cross-national ultimatum game experiments. Journal of Economic Behavior & Organization 64, 35 48.

Cooper, D.J., Feltovich, N., Roth, A.E., Zwick, R., 2003. Relative versus absolute speed of adjustment in strategic environments: Responder behavior in ultimatum games. Experimental Economics 6, 181 207.

Cooper, D.J., Kagel, J.H., 2016. Other regarding preferences: A selective survey of experimental results, in: Kagel, J.H., Roth, A.E. (Eds.), Handbook of Experimental Economics, Volume 2. Princeton: Princeton University Press, pp. 217 289.

Cueva, C., Iturbe-Ormaetxe, I., Mata-Pérez, E., Ponti, G., Sartarelli, M., Yu, H., Zhukova, V., 2016. Cognitive (ir)reflection: New experimental evidence. Journal of Behavioral and Experimental Economics 64, 81 93.

Eckel, C.C., Grossman, P.J., 2001. Chivalry and solidarity in ultimatum games. Economic Inquiry 39, 171 188.

Eckel, C.C., Grossman, P.J., 2002. Sex differences and statistical stereotyping in attitudes towards financial risks. Evolution and Human Behavior 23, 281 295.

Eckel, C.C., Grossman, P.J., 2008. Men, women and risk aversion: Experimental evidence, in: Plott, C., Smith, V. (Eds.), Handbook of Experimental Economics Results, Volume 1. New York: Elsevier, pp. 1061 1073.

Forsythe, R., Horowitz, J.L., Savin, N.E., Sefton, M., 1994. Fairness in simple bargaining experiments. Games and Economic Behavior 6, 347 369.

Gal, I., Baron, J., 1996. Understanding repeated simple choices. Thinking & Reasoning 2, 81 98.

Gale, J., Binmore, K., Samuelson, L., 1995. Learning to be imperfect: The ultimatum game. Games and Economic Behavior 8, 56 90.

Garcia-Gallego, A., Georgantzis, N., Jaramillo-Gutiérrez, A., 2012. Gender differences in ultimatum games: Despite rather than due to risk attitudes. Journal of Economic Behavior & Organization 83, 42 49.

Georganas, S., Healy, P.J., Weber, R.A., 2015. On the persistence of strategic sophistication. Journal of Economic Theory 159, 369 400.

Goodnow, J.J., 1955. Response sequences in a pair of two-choice probability situations. American Journal of Psychology 68, 624 630.

Grant, D.A., Hake, H.W., Hornseth, J.P., 1951. Acquisition and extinction of a verbal conditioned response with differing percentages of reinforcement. Journal of Experimental Psychology 42, 1 5.

Güth, W., Kocher, M.G., 2014. More than thirty years of ultimatum bargaining experiments: Motives, variations, and a survey of the recent literature. Journal of Economic Behavior & Organization 108, 396 409.

Güth, W., Schmittberger, R., Schwarze, B., 1982. An experimental analysis of ultimatum bargaining. Journal of Economic Behavior & Organization 3, 367 388.

Harrison, G.W., Lau, M.I., Rutström, E.E., 2007. Estimating risk attitudes in Denmark: A field experiment. Scandinavian Journal of Economics 109, 341 368.

Henrich, J., 2000. Does culture matter in economic behavior? Ultimatum game bargaining among the Machiguenga of the Peruvian Amazon. American Economic Review 90, 973 979.

Holt, C.A., Laury, S.K., 2002. Risk aversion and incentive effects. American Economic Review 92, 1644 1655.

Holt, C.A., Laury, S.K., 2014. Assessment and estimation of risk preferences, in: Machina, M.J., Viscusi, W.K. (Eds.), Handbook of the Economics of Risk and Uncertainty, Volume 1. North-Holland, Amsterdam, pp. 135 201.

Kim, D.G., Kim, H.C., 2021. Probability matching and strategic decision making.

Koehler, D.J., James, G., 2010. Probability matching and strategy availability. Memory & Cognition 38, 667 676.

Li, S., Qin, X., Houser, D., 2018. Revisiting gender differences in ultimatum bargaining: experimental evidence from the us and china. Journal of the Economic Science Association 4, 180 190.

Machina, M.J., Viscusi, W.K. (Eds.), 2014. Handbook of the Economics of Risk and Uncertainty, Volume 1. North-Holland, Amsterdam.

Neimark, E.D., Shuford, E.H., 1959. Comparison of predictions and estimates in a probability learning situation. Journal of Experimental Psychology 57, 294 298.

Noussair, C.N., Trautmann, S.T., van de Kuilen, G., 2014. Higher order risk attitudes, demographics, and financial decisions. Review of Economic Studies 81, 325 355.

Oechssler, J., Roider, A., Schmitz, P.W., 2009. Cognitive abilities and behavioral biases. Journal of Economic Behavior & Organization 72, 147 152.

Oosterbeek, H., Sloof, R., van de Kuilen, G., 2004. Cultural differences in ultimatum game experiments: Evidence from a meta-analysis. Experimental Economics 7, 171 188.

Roth, A.E., Erev, I., 1995. Learning in extensive-form games: Experimental data and simple dynamic models in the intermediate term. Games and Economic Behavior 8, 164 212.

Rubinstein, A., 2002. Irrational diversification in multiple decision problems. European Economic Review 46, 1369 1378.

Shanks, D.R., Tunney, R.J., McCarthy, J.D., 2002. A re-examination of probability matching and rational choice. Journal of Behavioral Decision Making 15, 233 250.

Solnick, S.J., 2001. Gender differences in the ultimatum game. Economic Inquiry 39, 189 200.

Thaler, R.H., 1988. The ultimatum game. Journal of Economic Perspectives 2, 195 206.

Thaler, R.H., 2016. Behavioral economics: Past, present, and future. American Economic Review 106, 1577 1600.

Vulkan, N., 2000. An economist's perspective on probability matching. Journal of economic surveys 14, 101 118.

West, R.F., Stanovich, K.E., 2003. Is probability matching smart? Associations between probabilistic choices and cognitive ability. Memory & Cognition 31, 243 251.

A Appendix: Experimental instructions

In this appendix we present the English translations of the (Chinese) instructions used in the Explicit treatment of the UG (subsection A.1), the Implicit treatment of the UG (subsection A.2), the Probability Matching test (subsection A.3), and the Risk Attitude test (subsection A.4).

A.1 Explicit treatment UG

Instructions

This is a **"money distribution with COMPUTER"** game between you and a **COMPUTER opponent**.

In this game, each round you need to decide to offer some "game points" from a pie which has 200 "game points" to your COMPUTER opponent and to keep the remaining "game points".

So you need to divide 200 game points into two parts in each round: One part for the COMPUTER opponent, and the rest for yourself.

After you make the decision, the COMPUTER opponent will choose to accept or reject your offer with a fixed probability based on how many game points you offer.

If the COMPUTER opponent chooses to accept, you will get the game points of your part. Otherwise, you will get nothing.

There are four options you can choose as follows:

Acceptance probabilities

Your offer (to computer player)	**50**	**100**	**150**	**200**
Acceptance probabilities of COMPUTER	44%	55%	70%	100%

About your payoffs:

After 200 rounds, we will **sum** your game points as the final payoff for this part of the experiment.

The first row of the chart is the game points that you can offer your COMPUTER opponent. There are four options here. If you decide to offer "50 game points" to your COMPUTER opponent, that means you want to keep the remaining "150 game points". If the COMPUTER opponent chooses to accept, you will get this "150 game points"' in this round. But if the COMPUTER opponent chooses to reject, you will get nothing in this round.

The second row of the chart shows the fixed probabilities of acceptance for each option. Those acceptance probabilities are given in advance and fixed throughout the whole game. For example, if you decide to offer your COMPUTER opponent "50 game points", there is a 44% chance that the COMPUTER opponent chooses to "accept" this option. It means you can get "150 game points" with a 44% chance.

In general, **the more you offer, the higher the acceptance probabilities by your COMPUTER opponent and the less game points you leave for yourself.**

In each round, please click to select the option you want to take, the COMPUTER will return a result for you according to the rules above.

A.2 Implicit treatment UG

Instructions

This is a **"money distribution with COMPUTE"** game between you and a **COMPUTER opponent**.

In this game, each round you need to decide to offer some "game points"' from a pie which has 200 "game points" to your COMPUTER opponent and to keep the remaining "game points"'.

So you need to divide 200 game points into two parts in each round: One part for the COMPUTER opponent, and the rest for yourself.

After you make the decision, the COMPUTER opponent will choose to accept or reject your offer with a fixed probability based on how many game points you offer.

If the COMPUTER opponent chooses to accept, you will get the game points of your part. Otherwise, you will get nothing.

There are four options you can choose as follows:

Acceptance probabilities

Your offer (to computer player)	**50**	**100**	**150**	**200**
Acceptance probabilities of COMPUTER	+	++	+ + +	+ + ++

About your payoffs:

After 200 rounds, we will **sum** your game points as the final payoff for this part of the experiment.

The first row of the chart is the game points that you can offer your COMPUTER opponent. There are four options here. If you decide to offer "50 game points” to your COMPUTER opponent, that means you want to keep the remaining “150 game points”. If your COMPUTER opponent chooses to accept, you will get this “150 game points” in this round. But if your COMPUTER opponent chooses to reject, you will get nothing in this round.

The second row of the chart shows the fixed probabilities of acceptance for each option. **Those acceptance probabilities are given in advance and fixed throughout the whole game**. **Please note that the more “+″ here only means the more likely the option is to be accepted, but it does not indicate any precise relationship between these probabilities.**

For example, option “100” with “++” acceptance probability is not twice as likely to be accepted as option “50” with “+” acceptance probability.

In general, **the more you offer, the higher the acceptance probabilities by your COMPUTER opponent and the less game points you leave for yourself.**

In each round, please click to select the option you want to take, the COMPUTER will return a result for you according to the rules above.

A.3 PM

Instructions

This is a 10-round “guessing number” game.

In each round of the game, the computer will show one number out of **0** and **1** to you. That means either **0** or **1** will be displayed after you guess a number.

The probability of showing **0** is 30%, the probability of showing **1** is 70%.

What you need to do is guess which number will be shown:

Click to select the number you guess and press the “next” button.

The computer will return to you a result in this round according to the above rules.

About your payoffs:

In the game, we will give you a certain number of “game points” as your payoffs based on the result of your choices. And after the end of the experiment, your “game points” will be converted into monetary rewards.

In each round, if the number you guess is the **same** as the computer shows, you will get **100 “game points”**. But if what you guess is **different** from what the computer shows, you will get **0 “game points”**. That means you get nothing if you guess the wrong number.

In game “guessing number”, the total payoff of this part is **the sum of all the payoffs for 10 rounds**.

A.4 RA

Instructions

This is a “Ten Paired Lottery-Choice Decisions” game.

There are ten different situations in this game. Each of them contains a paired Lottery-Choice decision called “GAME A” and “GAME B”. For each of the situations, you need to **make a decision to choose one of the two games: “GAME A” or “GAME B”**. Once selected, it means you want play the game you choose under this situation. The computer will **randomly select one of ten situations** as the **“paid situation”**. In this situation, “GAME A” will be played if you choose “GAME A”, and “GAME B” will be played if you choose “GAME B”. You will be paid according to the result.

"Ten Paired Lottery-Choice Decisions" game

Make a decision for every situation	GAME A	GAME B
Situation 1	1/10 of ¥2.00, 9/10 of ¥1.60	1/10 of ¥3.85, 9/10 of ¥0.10
Situation 2	2/10 of ¥2.00, 8/10 of ¥1.60	2/10 of ¥3.85, 8/10 of ¥0.10
Situation 3	3/10 of ¥2.00, 7/10 of ¥1.60	3/10 of ¥3.85, 7/10 of ¥0.10
Situation 4	4/10 of ¥2.00, 6/10 of ¥1.60	4/10 of ¥3.85, 6/10 of ¥0.10
Situation 5	5/10 of ¥2.00, 5/10 of ¥1.60	5/10 of ¥3.85, 5/10 of ¥0.10
Situation 6	6/10 of ¥2.00, 4/10 of ¥1.60	6/10 of ¥3.85, 4/10 of ¥0.10
Situation 7	7/10 of ¥2.00, 3/10 of ¥1.60	7/10 of ¥3.85, 3/10 of ¥0.10
Situation 8	8/10 of ¥2.00, 2/10 of ¥1.60	8/10 of ¥3.85, 2/10 of ¥0.10
Situation 9	9/10 of ¥2.00, 1/10 of ¥1.60	9/10 of ¥3.85, 1/10 of ¥0.10
Situation 10	10/10 of ¥2.00, 0/10 of ¥1.60	10/10 of ¥3.85, 0/10 of ¥0.10

B Appendix: Variable definitions

This appendix presents a complete list of all the variables used in the paper in alphabetical order.

Age: The age as reported by the player, measured in years.

Age difference: As the players' age was (unsurprisingly) highly correlated with the grade they were studying for, we constructed the variable *Age difference* by taking for each player the difference between their age and the average age of players at the grade at which they were currently studying.

Degree: This dummy variable indicates whether the player was studying for a degree in the field of Economics, Finance or Business Studies in a broad sense (value of 1), or not (value of 0).

Gender: A dummy variable equal to 1 (male) or 0 (female).

Grade: This is a variable indicating the grade at which the players were currently studying: 0 'Freshman', 1 'Sophomore', 2 'Junior', 3 'Senior', 4 'Master's First grade', 5 'Master's Second grade', 6 'Master's Third grade', 7 'PhD First grade', 8 'PhD Second grade', 9 'PhD Third grade', 10 'Other'.

Order tasks: The RA test was always done at the end of a session. But the PM test was sometimes done before and sometimes after the UG. *Order tasks* is a dummy variable, equal to 1 if the order was PM-UG-RA, and equal to 0 if the order was UG-PM-RA.

Participations: This variable indicates the number of previous participations in experiments.

PM pass: This is another dummy variable, equal to 1 (Pass) for players making exclusively optimal choices, thereby passing the Probability Matching (PM) test, and equal to 0 (Fail) for all other players.

RA pass: This is a dummy variable, equal to 1 (Pass) if the RA test indicates that for this player the optimal UG offer is equal to the minimum offer of 50. This applies to players with a regular choice pattern who are risk-seeking, risk-neutral, or mildly risk-averse (i.e. players with 0 up to and including 5 safe lottery choices before switching to the risky lotteries). The dummy is equal to 0 (Fail) if we have no such evidence. This applies to players with a later switching point or with seriously irregular choice patterns.

Treatment: There are two treatments. In the Explicit treatment all acceptance probabilities of the UG are given upfront to the subjects. In the Implicit treatment these probabilities are still fixed and known to be fixed, but they are not given upfront to the subjects.

UG offer: This is the main dependent variable. For each individual subject we compute the average offer over their history of play or a subset thereof. Most of the results in our analysis concern the entire history of 200 periods. In Appendix D we also consider a number of alternative dependent variables, computing the average UG offer instead over different subhistories of play: periods 1-10, 1-50, 1-100, 101-200, 151-200 and 191-200, respectively.

University: The Explicit treatment sessions were conducted at Beijing Normal University (BNU) and Zhejiang University (ZJU), whereas the Implicit sessions were at BNU and Zhejiang Gongshang University (ZJSU). Therefore this dummy variable was equal to 0 for BNU players (in either treatment), and 1 for ZJU or ZJSU players (in Explicit and Implicit treatments, respectively).

C Appendix: UG offers for various groups

In this appendix we show some graphs illustrating how the UG offers relate to the two focus variables (*RA pass* and *PM pass*) and *Gender*. We first consider the average UG offers and then look at the distributions over the various UG offer levels.

Figure 8 presents the average offers in the Implicit and Explicit treatments for players who passed the RA test and those who failed this test, with the inset showing again how these averages evolve during the 200 periods. Figure 9 shows the same for players who passed or failed the PM test, and Figure 10 does so for male and female players in both treatments. Note that while there are some differences in the levels (Implicit v. Explicit, as well as with respect to the variables considered here, more pronounced in the Explicit treatment), all these averages evolve over time in a very similar way.

Figures 11, 12 and 13 present the distribution of the individual averages over the 200 periods in the two treatments, distinguishing those passing or failing the RA test or PM test as well as male and female players.[29] In all these three figures we see similar differences between the Implicit and Explicit treatment as in Figure 6, with the Explicit distributions shifted somewhat to the left and with a peak for the first bin (labeled "50"), whereas this peak is absent in the Implicit distribution.

[29]As in Figure 7, bin "50" = [50, 55), bin "55" = [55, 60), . . . , . . . , bin "150+" = [150, 200].

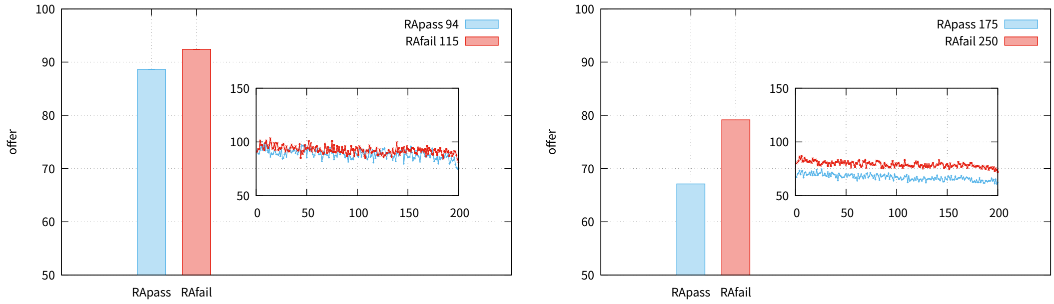


Figure 8: Average offers in Implicit and Explicit treatment, RA test

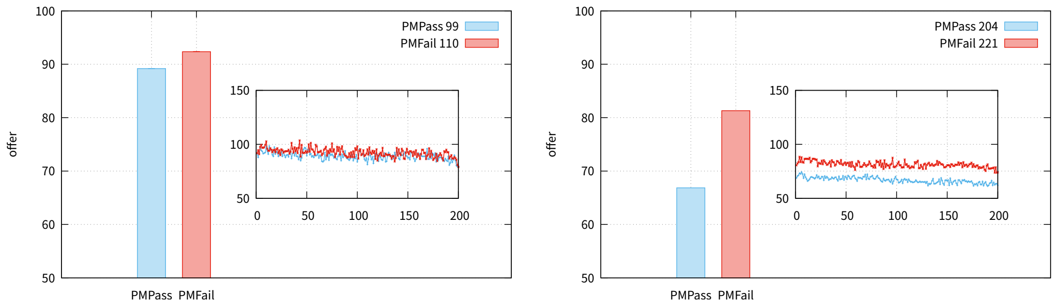


Figure 9: Average offers in Implicit and Explicit treatment, PM test

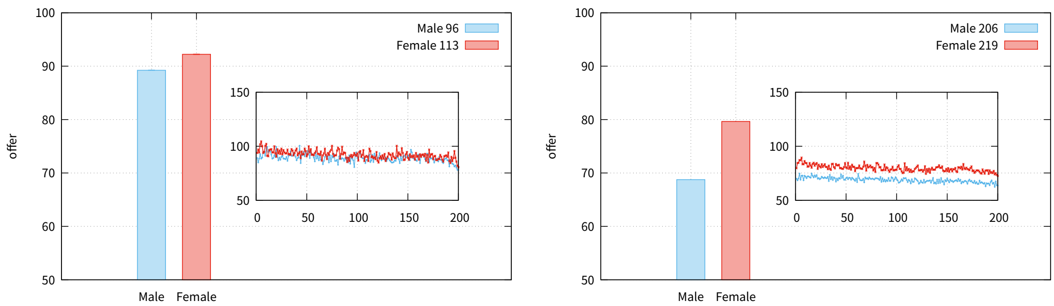


Figure 10: Average offers in Implicit and Explicit treatment, Gender

When we consider the effect of these variables within each of these charts (RA and PM test and gender), we see two effects. First, in each of the charts (for both Implicit and Explicit treatment) we see a relatively modest shift in the distribution to the left for those passing the RA test or PM test and for male players. Second, in the Explicit treatment we see an enormous jump in the frequency for the first bin, more than doubling, from about 20% to about 40% for those passing the RA or PM test and for male players.

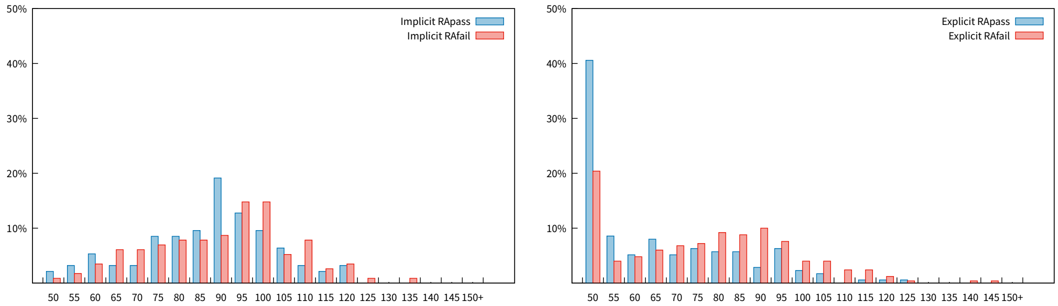


Figure 11: Distribution of offers in Implicit and Explicit treatment, RA test

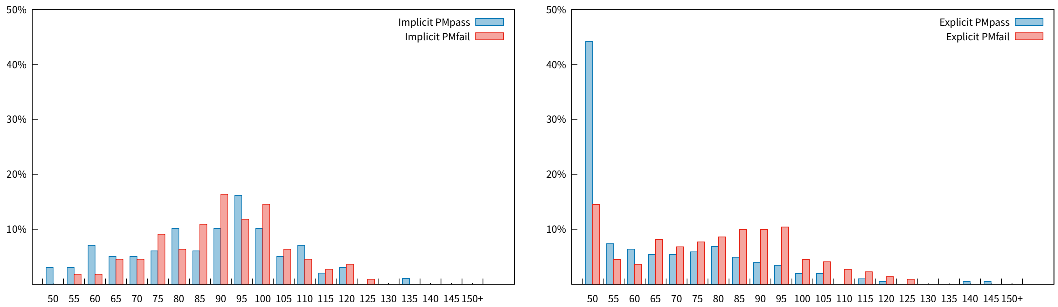


Figure 12: Distribution of offers in Implicit and Explicit treatment, PM test

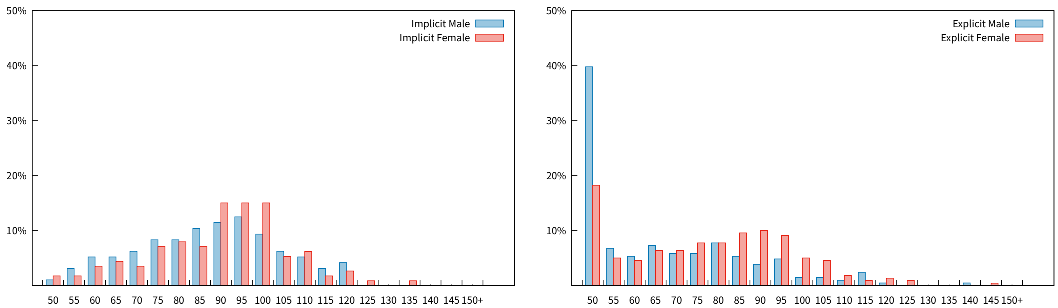


Figure 13: Distribution of offers in Implicit and Explicit treatment, Gender

D Appendix: Alternative dependent variables

Tables 4 and 5 show the Tobit regression results for alternative specifications of the dependent variable in the Implicit and Explicit treatment respectively. More precisely, we consider alternative intervals, either at the beginning or at the end of the history of 200 periods, to check for the robustness of the results for the parsimonious specifications reported in Table 3 in the main text.

Table 4 gives the results for the Implicit treatment. Initially (in periods 1-10) only *Gender* seems to matter (-7.4 at 5%), not too different from what we saw in the Explicit treatment. But this quickly seems to be overwhelmed by other effects (incl. possibly confusion), disappearing completely. Similarly, *Age difference* seems to be relevant in the first half of the history only. We note that *RA pass* becomes more significant over time in the Implicit treatment. In periods 191-200 players for whom the optimal UG offer is 50 do indeed offer clearly less than the other players (-8.1 at 5%), approaching the magnitude and significance we see in the Explicit treatment.

As we see in Table 5, in the Explicit treatment the results are relatively robust across the 200 periods. In particular, we see the same kind of relevance (in terms of the magnitude of the coefficient and statistical significance) of the most important variables highlighted in our analysis in the main text: *RA pass*, *PM pass* and *Gender*. The main aspect that changes over time is that UG offers tend to decrease somewhat, and this seems mainly reflected in the reduced size of the intercept, from 101.1 in periods 1-10 to 84.5 in periods 191-200, without however showing a clear change in the relative importance and significance of the main variables. But there are some small differences in a couple of the other variables. Initially, players at ZJU tend to make substantially lower offers than those at BNU (-11.2, significant at 1%), but this effect quickly fades away. Similarly, players who are older than their peers tend to make higher UG offers (about 2-3/year at 1% significance), but this is the case only in the first half of the 200 periods.

Treatment	Implicit					
	Dependent variable: individual UG offers averaged over					
	(1)	(2)	(3)	(4)	(5)	(6)
Periods	1-10	1-50	1-100	101-200	151-200	191-200
Intercept	97.8927***	100.6591***	99.5150***	93.0965***	94.3122***	91.2119***
	(2.0485)	*(2.4243)*	*(2.3526)*	*(2.3145)*	*(3.0145)*	*(3.5387)*
RA pass		−4.3449*	−4.3635*		−5.4529*	−8.1024**
		(2.3047)	*(2.2364)*		*(2.8933)*	*(3.4072)*
PM pass						
Order tasks						
University						
Grade (level)		−1.6965***	−1.6481***	−1.3559**	−1.2710*	−1.4654*
		(0.5995)	*(0.5817)*	*(0.6793)*	*(0.7547)*	*(0.8870)*
Age difference		2.0527*	2.0188*			
		(1.2158)	*(1.1800)*			
Participations						
Degree (field)						
Gender	−7.4360**					
	(3.0282)					
Observations	209	209	209	209	209	209
LR χ^2	5.95	12.55	12.82	3.95	5.48	7.25
(*degr.freed.*)	*(1)*	*(3)*	*(3)*	*(1)*	*(2)*	*(2)*
(*p-value*)	*(0.0147)*	*(0.0057)*	*(0.0051)*	*(0.0470)*	*(0.0645)*	*(0.0266)*
F for restrictions	0.71	0.54	0.91	0.95	0.77	0.87
(*degr.freed.*)	*(8, 200)*	*(6, 200)*	*(6, 200)*	*(8, 200)*	*(7, 200)*	*(7, 200)*
(*p-value*)	*(0.6817)*	*(0.7755)*	*(0.4868)*	*(0.4726)*	*(0.6137)*	*(0.5313)*

Notes: All variables are described in the main text and in Appendix B. For each variable we report the estimated coefficient with the standard error in brackets. For each column, the number of observations corresponds to the number of individual players. For each column we report the Likelihood Ratio (LR) χ^2-test that at least one of the predictors' regression coefficients is not equal to zero (with the degrees of freedom and p-value in brackets). We also report the F-test for the joint restrictions, starting from the complete specification analogous to that reported in the main text, with the corresponding degrees of freedom and p-value in brackets. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$

Table 4: Tobit regression additional results Implicit treatment: alternative dependent variables

Treatment	Explicit					
	Dependent variable: individual UG offers averaged over					
	(1)	(2)	(3)	(4)	(5)	(6)
Periods	1-10	1-50	1-100	101-200	151-200	191-200
Intercept	101.0980***	91.9925***	90.9702***	89.1372***	88.4348***	84.4602***
	(3.2710)	*(2.6047)*	*(2.2707)*	*(2.5152)*	*(2.6692)*	*(3.0068)*
RA pass	−13.4166***	−9.2648***	−10.5629***	−12.0033***	−12.5956***	−12.8663***
	(3.0817)	*(2.6047)*	*(2.3928)*	*(2.6698)*	*(2.8506)*	*(3.2508)*
PM pass	−15.1310***	−13.8446***	−13.6554***	−19.0474***	−20.0587***	−19.7767***
	(3.0017)	*(2.5480)*	*(2.3178)*	*(2.6375)*	*(2.8166)*	*(3.2085)*
Order tasks	−5.4183*	−4.7912*		−5.1840**	−5.1507*	−5.4381*
	(2.9269)	*(2.4861)*		*(2.5645)*	*(2.7312)*	*(3.1038)*
University	−11.1507***	−4.8150*	−3.9627*			
	(2.9598)	*(2.5036)*	*(2.2979)*			
Grade (level)						
Age difference	3.3378***	2.1847**	2.0779**			
	(1.2129)	*(1.0343)*	*(0.9486)*			
Participations						
Degree (field)						
Gender	−14.7335***	−6.0384**	−9.8926***	−9.3403***	−10.2439***	−11.3028***
	(2.9912)	*(2.5357)*	*(2.3258)*	*(2.6044)*	*(2.7762)*	*(3.1648)*
Observations	425	425	425	425	425	425
LR χ^2	103.37	68.01	95.29	106.15	104.99	86.22
(*degr.freed.*)	*(6)*	*(6)*	*(5)*	*(4)*	*(4)*	*(4)*
(*p-value*)	*(0.0000)*	*(0.0000)*	*(0.0000)*	*(0.0000)*	*(0.0000)*	*(0.0000)*
F for restrictions	0.35	0.42	0.99	1.32	0.86	0.70
(*degr.freed.*)	*(3, 416)*	*(3, 416)*	*(4, 416)*	*(5, 416)*	*(5, 416)*	*(5, 416)*
(*p-value*)	*(0.7913)*	*(0.7414)*	*(0.4127)*	*(0.2561)*	*(0.5104)*	*(0.6262)*

Notes: All variables are described in the main text and in Appendix B. For each variable we report the estimated coefficient with the standard error in brackets. For each column, the number of observations corresponds to the number of individual players. For each column we report the Likelihood Ratio (LR) χ^2-test that at least one of the predictors' regression coefficients is not equal to zero (with the degrees of freedom and p-value in brackets). We also report the F-test for the joint restrictions, starting from the complete specification analogous to that reported in the main text, with the corresponding degrees of freedom and p-value in brackets. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$

Table 5: Tobit regression additional results Explicit treatment: alternative dependent variables

E Appendix: Dropping 'confused' players

In Table 6 we show the Tobit regression results for the same specifications as in Table 3 in the main text, i.e., with focus variables only in columns (1) and (4), a complete specification in columns (2) and (5) and a more parsimonious one in columns (3) and (6), with the first three columns being for the Implicit treatment and the last three for the Explicit treatment. In this table we excluded the players who rarely chose an offer of 50 (strictly less than 20 out of 200 times) although their RA test indicated that this was their optimal UG offer. As we see, this left us with 198 players in the Implicit treatment and 423 in the Explicit treatment (out of 209 and 425 for the Implicit and Explicit treatment respectively).

For the Implicit treatment, we see that all three regression specifications are now significant. In the specification with only the focus variables in column (1), both the RA and PM test are significant and with the expected sign. Adding the control variables, in the complete specification in column (2) the PM test loses its significance whereas the grade at which the players study is weakly significant. In the more parsimonious specification in column (3) both RA and PM test are (weakly) significant as is the players' grade.

For the Explicit treatment the table merely confirms the results reported in the main text when all observations were included. Just as with the regression results including all observations, we see in Table 6 that the coefficients and significances are reasonably consistent across these specifications for both treatments.

	Dependent variable: individual UG offers averaged over periods 1-200					
	(1)	(2)	(3)	(4)	(5)	(6)
Treatment	Implicit	Implicit	Implicit	Explicit	Explicit	Explicit
Intercept	94.4291***	97.1501***	98.9937***	84.4085***	91.6963***	89.7149***
	(1.8555)	*(3.3809)*	*(2.5703)*	*(1.6530)*	*(3.0106)*	*(2.1012)*
RA pass	−5.1037**	−5.7641**	−6.0618**	−11.7476***	−11.0285***	−10.3427***
	(2.3558)	*(2.3470)*	*(2.3496)*	*(2.2485)*	*(2.2267)*	*(2.2073)*
PM pass	−4.9056**	-3.8359	−4.3620*	−15.9864***	−14.6860***	−15.0145***
	(2.3368)	*(2.3755)*	*(2.3101)*	*(2.2107)*	*(2.1743)*	*(2.1786)*
Order tasks		2.3757			−3.8605*	−3.7566*
		(2.3405)			*(2.1322)*	*(2.1258)*
University		-0.6767			-2.4176	
		(2.7858)			*(2.2355)*	
Grade (level)		−1.3989**	−1.5583**		0.1807	
		(0.6315)	*(0.6174)*		*(0.4920)*	
Age difference		1.5999			1.7371*	1.6833*
		(1.2546)			*(0.8850)*	*(0.8838)*
Participations		1.7094			-1.9638	
		(1.5961)			*(1.5773)*	
Degree (field)		1.6010			-0.5333	
		(2.7121)			*(3.9368)*	
Gender		-2.5411			−9.1684***	−9.2038***
		(2.3530)			*(2.1865)*	*(2.1705)*
Observations	198	198	198	423	423	423
LR χ^2	8.84	20.54	15.11	86.93	114.77	110.62
(degr.freed.)	*(2)*	*(9)*	*(3)*	*(2)*	*(9)*	*(5)*
(p-value)	*(0.0120)*	*(0.0149)*	*(0.0017)*	*(0.0000)*	*(0.0000)*	*(0.0000)*
F for restrictions	1.72	n.a.	0.92	4.09	n.a.	1.04
(degr.freed.)	*(7, 189)*	*n.a.*	*(6, 189)*	*(7, 414)*	*n.a.*	*(4, 414)*
(p-value)	*(0.1063)*	*n.a.*	*(0.4842)*	*(0.0002)*	*n.a.*	*(0.3843)*

Notes: All variables are described in the main text and in Appendix B. For each variable we report the estimated coefficient with the standard error in brackets. For each column, the number of observations corresponds to the number of individual players. For each column we report the Likelihood Ratio (LR) χ^2-test that at least one of the predictors' regression coefficients is not equal to zero (with the degrees of freedom and p-value in brackets). In columns (1), (3), (4) and (6) we also report the F-test for the joint restrictions, starting from the complete specifications in columns (2) and (5), with the corresponding degrees of freedom and p-value in brackets. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$

Table 6: Tobit regression dropping some 'confused' players

School of Economics and Finance

Queen Mary
University of London

School of Economics and Finance Queen Mary University of London
Mile End Road
London E1 4NS
Tel: +44 (0)20 7882 7356
Fax: +44 (0)20 8983 3580
Web: www.econ.qmul.ac.uk/research/workingpapers/